

ALMOST SURE CONVERGENCE IN QUANTUM SPIN GLASSES

DAVID BUZINSKI AND ELIZABETH MECKES

ABSTRACT. Recently, Keating, Linden, and Wells [7] showed that the density of states measure of a nearest-neighbor quantum spin glass model is approximately Gaussian when the number of particles is large. The density of states measure is the ensemble average of the empirical spectral measure of a random matrix; in this paper, we use concentration of measure and entropy techniques together with the result of [7] to show that in fact, the empirical spectral measure of such a random matrix is almost surely approximately Gaussian itself, with no ensemble averaging. We also extend this result to a spherical quantum spin glass model and to the more general coupling geometries investigated by Erdős and Schröder.

1. INTRODUCTION AND STATEMENTS OF RESULTS

In the recent paper [7], Keating, Linden and Wells show that the density of states measure of a quantum spin glass with nearest neighbor interactions and Gaussian coupling coefficients is approximately Gaussian, as the number of particles tends to infinity. More specifically, they considered the following random matrix model for the Hamiltonian of a quantum spin glass: let $\{Z_{a,b,j}\}_{\substack{1 \leq a,b \leq 3 \\ 1 \leq j \leq n}}$ be independent standard Gaussian random variables, and define the $2^n \times 2^n$ random matrix H_n by

$$(1) \quad H_n := \frac{1}{\sqrt{9n}} \sum_{j=1}^n \sum_{a,b=1}^3 Z_{a,b,j} \sigma_j^{(a)} \sigma_{j+1}^{(b)},$$

where for $1 \leq a \leq 3$,

$$\sigma_j^{(a)} := I_n^{\otimes(j-1)} \otimes \sigma^{(a)} \otimes I_2^{\otimes(n-j)},$$

with I_2 denoting the 2×2 identity matrix, $\sigma^{(a)}$ denoting the 2×2 non-trivial Pauli matrices

$$\sigma^{(1)} := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma^{(2)} := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \sigma^{(3)} := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

and the labeling cyclic so that $\sigma_{n+1}^{(b)} := \sigma_1^{(b)}$.

The density of states measure μ_n^{DOS} for the system is the ensemble average of the spectral measure of H_n ; that is, if $\{\lambda_j\}_{1 \leq j \leq 2^n}$ are the (necessarily real) eigenvalues of H_n , then for $A \subseteq \mathbb{R}$,

$$\mu_n^{DOS}(A) = \frac{1}{2^n} \mathbb{E} |\{j : \lambda_j \in A\}|.$$

In other words, $\mu_n^{DOS}(A)$ is the expected proportion of the eigenvalues of H_n lying in the set A . The main result of [7] is that μ_n^{DOS} converges weakly to Gaussian, as $n \rightarrow \infty$. The authors go on to consider more general collections of (still independent) coupling coefficients, and more general coupling geometries than that of nearest-neighbor interactions. In more recent work, Erdős and Schröder [3] have considered still more general coupling geometries, and found a sharp transition in the limiting behavior of the density of states measure

depending on the size of the maximum degree of the underlying graph, relative to its number of edges.

The purpose of this paper is to move from convergence in expectation of the spectral measure of H_n to the considerably stronger notion of almost sure convergence. As observed in [7], the extent to which ensemble averages actually manage to describe features of individual systems is not always clear, but is a crucial issue if one is to make meaningful use of random matrix models. The following result shows that the Gaussian behavior exhibited by the average spectral measure is indeed the typical behavior of the empirical spectral measures.

Theorem 1. *Let μ_n be the spectral measure of H_n and let γ denote the standard Gaussian distribution. There are universal constants C , C' and c such that*

$$\begin{aligned} (a) \quad & \mathbb{E}d_{BL}(\mu_n, \gamma) \leq \frac{C}{n^{1/6}}; \\ (b) \quad & \mathbb{P} \left[d_{BL}(\mu_n, \gamma) \geq \frac{C}{n^{1/6}} + t \right] \leq C e^{-c n t^2}; \end{aligned}$$

and

(c) with probability 1, for all sufficiently large n ,

$$d_{BL}(\mu_n, \gamma) \leq \frac{C'}{n^{1/6}}.$$

Here $d_{BL}(\mu, \nu)$ denotes the bounded-Lipschitz distance between probability measures, which metrizes the topology of weak convergence.

In the paper [6], Keating, Linden and Wells took a different approach to understanding the behavior of the spectral measures of individual Hamiltonians; rather than consider a random matrix model, they took the coefficients in (1) to be deterministic, subject to a normalization and boundedness condition, and showed that in that case, the non-random spectral measures converged weakly to Gaussian. It should be possible to take that result as a starting point in order to obtain almost sure convergence in the Gaussian model, although there are various technical challenges. We instead take a rather different approach, combining convergence in expectation with various probabilistic techniques.

The organization of this paper is as follows. In Section 2, the pointwise estimate of [7] on the difference between characteristic functions of μ_n^{DOS} and γ is parlayed into an estimate on $d_{BL}(\mu_n^{DOS}, \gamma)$ using Fourier analysis. In Section 3 the Gaussian concentration of measure phenomenon is used to show that the random variable $d_{BL}(\mu_n, \gamma)$ is strongly concentrated at its mean. Then, the expected distance between μ_n and its average μ_n^{DOS} is estimated; this is done using a combination of applications of Gaussian concentration of measure, entropy methods, and approximation theory, via a similar approach to the one taken by the second author and M. Meckes in [9]. The almost-sure convergence rate given in part (c) is an immediate consequence of part (b) and the first Borel-Cantelli lemma, and is therefore not discussed further. In Section 4, we consider a modification of the random matrix model above, in which the coefficients in H_n are not independent but are drawn uniformly from the $9n$ -dimensional sphere, and show that the empirical spectral measure is almost surely approximately Gaussian in that setting as well. Finally, Section 5 offers some remarks on extensions of Theorem 1 to further related ensembles.

Notation and Conventions. Let the random matrix H_n be defined as above, with eigenvalues $\lambda_1 < \dots < \lambda_{2^n}$. The empirical spectral measure μ_n of H_n is defined by

$$\mu_n := 2^{-n} \sum_{j=1}^{2^n} \delta_{\lambda_j};$$

that is, μ_n is the random probability measure putting equal mass at each eigenvalue of H_n . Its ensemble average $\mathbb{E}\mu_n$ is denoted μ_n^{DOS} and is called the density of states measure.

For probability measures μ and ν on \mathbb{R} , the bounded-Lipschitz distance $d_{BL}(\mu, \nu)$ from μ to ν is defined by

$$d_{BL}(\mu, \nu) := \sup \left\{ \left| \int f d\mu - \int f d\nu \right| : \|f\|_{BL} \leq 1 \right\}$$

where $\|f\|_{BL}$ denotes the bounded-Lipschitz norm of f , defined by

$$\|f\|_{BL} := \|f\|_{\infty} + |f|_L,$$

with $|f|_L$ denoting the Lipschitz constant of f . The bounded-Lipschitz distance metrizes weak convergence of probability measures.

If the test functions are required only to be Lipschitz and not necessarily bounded, one gets instead the L_1 -Kantorovich distance $W_1(\mu, \nu)$:

$$W_1(\mu, \nu) := \sup \left\{ \left| \int f d\mu - \int f d\nu \right| : |f|_L \leq 1 \right\}.$$

Clearly, $d_{BL}(\mu, \nu) \leq W_1(\mu, \nu)$.

It is the Kantorovich–Rubinstein theorem that W_1 is also given by

$$W_1(\mu, \nu) = \inf_{\pi} \int |x - y| d\pi(x, y),$$

where the infimum is taken over all couplings π of the measures μ and ν . It is for this reason that W_1 is also called the L_1 -coupling distance.

Finally, symbols such as C, c will denote universal constants independent of all parameters, which may change in value from one appearance to the next.

2. GAUSSIAN DENSITY OF STATES IN THE KANTOROVICH DISTANCE

The crucial ingredient in the estimation of $d_{BL}(\mu_n^{DOS}, \gamma)$ is the following pointwise bound from [7] on the difference between the corresponding characteristic functions.

Theorem 2 (Keating, Linden, and Wells). *Let μ_n^{DOS} and γ be as above; denote the characteristic function of μ_n^{DOS} by ψ_n and the characteristic function of γ by φ . There is a constant C independent of n such that for all ξ ,*

$$(2) \quad |\psi_n(\xi) - \varphi(\xi)| \leq \frac{C\xi^2}{\sqrt{n}}.$$

The main result of this section is the following.

Theorem 3. *Let μ_n^{DOS} and γ be as above. There is a constant c such that*

$$d_{BL}(\mu_n^{DOS}, \gamma) \leq \frac{c}{n^{1/4}}.$$

Note that by definition of the bounded-Lipschitz distance there is no loss in restricting to the case $f(0) = 0$, which we do for the remainder of the proof.

The first step is to make a truncation argument to further restrict the class of test functions considered. Given a function $f : \mathbb{R} \rightarrow \mathbb{R}$ with $\|f\|_{BL} \leq 1$ and $f(0) = 0$, and given $R > 0$, define the truncation f_R by

$$(3) \quad f_R(x) = \begin{cases} f(x), & |x| \leq R; \\ f(R) + [\operatorname{sgn}(f(R))](R - x), & R < x < R + |f(R)|; \\ f(-R) + [\operatorname{sgn}(f(-R))](x + R), & -|f(-R)| - R < x < -R; \\ 0, & x \leq -|f(-R)| - R \text{ or } x \geq R + |f(R)|. \end{cases}$$

Then $\|f_R\|_{BL} \leq 1$ and f_R is supported on $[-2R, 2R]$.

Lemma 4.

(a) For any $t > 0$,

$$\mu_n^{DOS}(\{x : |x| > t\}) \leq \frac{c}{t^2}.$$

(b) For $f : \mathbb{R} \rightarrow \mathbb{R}$ with $\|f\|_{BL} \leq 1$,

$$\left| \int (f - f_R) d\mu_n^{DOS} \right| \leq \frac{c}{R^2}.$$

Proof. A straightforward Fubini's theorem argument (see, e.g., Section 26 of [2]) gives that

$$\mu_n^{DOS}(\{x : |x| > t\}) \leq \frac{t}{2} \int_{-\frac{2}{t}}^{\frac{2}{t}} (1 - \psi_n(\xi)) d\xi,$$

where as before $\psi_n(\xi)$ denotes the characteristic function of μ_n^{DOS} . Adding and subtracting the characteristic function $\varphi(\xi)$ of the standard Gaussian distribution and using (2) then gives

$$\frac{t}{2} \int_{-\frac{2}{t}}^{\frac{2}{t}} (1 - \psi_n(\xi)) d\xi = \frac{t}{2} \int_{-\frac{2}{t}}^{\frac{2}{t}} (1 - \varphi(\xi)) d\xi + \frac{t}{2} \int_{-\frac{2}{t}}^{\frac{2}{t}} (\varphi(\xi) - \psi_n(\xi)) d\xi \leq \frac{c}{t^2} \left(1 + \frac{1}{\sqrt{n}} \right).$$

For part (b), note that by construction, $|f(x) - f_R(x)| \leq 1$; moreover, $f(x) = f_R(x)$ for $|x| \leq R$, so that

$$\left| \int f d\mu_n^{DOS} - \int f_R d\mu_n^{DOS} \right| \leq \int_{|x| > R} d\mu_n^{DOS}.$$

Part (a) is now immediate from part (b). □

Now let $f : \mathbb{R} \rightarrow \mathbb{R}$ have $\|f\|_{BL} \leq 1$ and $\operatorname{supp}(f) \subseteq [-2R, 2R]$. The next step in the proof of Theorem 3 is to approximate f by

$$f_\lambda := f * K_\lambda,$$

where K_λ is the Féjer kernel

$$K_\lambda(x) = \frac{1}{2\pi} \int_{-\lambda}^{\lambda} \left(1 - \frac{|\xi|}{\lambda} \right) e^{i\xi x} d\xi = \frac{\lambda}{2\pi} \left(\frac{\sin(\lambda x/2)}{\lambda x/2} \right)^2.$$

For f as above, one can approximate in the supremum norm, as follows.

Lemma 5. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ have $\|f\|_{BL} \leq 1$ and $\text{supp}(f) \subseteq [-2R, 2R]$. Then*

$$|f(x) - f_\lambda(x)| \leq \frac{8 \log(\lambda) + 8 \log(2R) + 6}{\pi \lambda}.$$

Proof. By definition of f_λ (using the second form of the Féjer kernel) and the fact that $\int_{-\infty}^{\infty} K_\lambda(y) dy = 1$,

$$|f(x) - f_\lambda(x)| = \frac{1}{\pi} \left| \int_{-\infty}^{\infty} \left(f(x) - f\left(x - \frac{2y}{\lambda}\right) \right) \left(\frac{\sin(y)}{y} \right)^2 dy \right|.$$

Now, if $x \in [-2R, 2R]$, then using the fact that f is supported on $[-2R, 2R]$ and is 1-Lipschitz yields

$$\begin{aligned} \pi |f(x) - f_\lambda(x)| &\leq |f(x)| \int_{-\infty}^{\frac{\lambda}{2}(x-2R)} \left| \frac{\sin(y)}{y} \right|^2 dy + \int_{\frac{\lambda}{2}(x-2R)}^{-1} \left| \frac{2 \sin^2(y)}{\lambda y} \right| dy + \int_{-1}^1 \left| \frac{2 \sin^2(y)}{\lambda y} \right| dy \\ &\quad + \int_1^{\frac{\lambda}{2}(2R-x)} \left| \frac{2 \sin^2(y)}{\lambda y} \right| dy + |f(x)| \int_{\frac{\lambda}{2}(2R-x)}^{\infty} \left| \frac{\sin(y)}{y} \right|^2 dy \\ &=: I + II + III + IV + V. \end{aligned}$$

Since $f(2R) = 0$,

$$I \leq |f(x) - f(2R)| \int_{-\infty}^{\frac{\lambda}{2}(x-2R)} \frac{1}{y^2} dy = \frac{2|f(x) - f(2R)|}{\lambda(2R-x)} \leq \frac{2}{\lambda};$$

V is handled the same way. Next, since $|x| \leq 2R$,

$$II \leq \int_{\frac{\lambda}{2}(x-2R)}^{-1} \frac{2}{\lambda|y|} dy = \frac{2}{\lambda} \log \left(\frac{\lambda}{2} (2R-x) \right) \leq \frac{4(\log(\lambda) + \log(2R))}{\lambda};$$

IV is the same. Finally, using the bound $\left| \frac{\sin(y)}{y} \right| \leq 1$ gives

$$III \leq \int_{-1}^1 \frac{2|y|}{\lambda} dy = \frac{2}{\lambda}.$$

If $x > 2R$, then by the concavity of the logarithm,

$$\begin{aligned} \pi |f(x) - f_\lambda(x)| &\leq \int_{\frac{\lambda}{2}(x-2R)}^{\frac{\lambda}{2}(x+2R)} \left| \frac{2 \sin^2(y)}{\lambda y} \right| dy \\ &\leq \frac{2}{\lambda} \left[\log \left(\frac{\lambda}{2} (x+2R) \right) - \log \left(\frac{\lambda}{2} (x-2R) \right) \right] \leq \frac{2}{\lambda} \left(\frac{4R}{x} \right) \leq \frac{4}{\lambda}. \end{aligned}$$

The case $x < -2R$ is the same. □

The following technical lemma is needed in order to compare $\int f_\lambda d\mu_n^{DOS}$ to $\int f_\lambda d\gamma$.

Lemma 6. *Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be such that $|g(\xi)| \leq \min \left\{ \frac{2}{|\xi|}, \frac{C|\xi|}{\sqrt{n}} \right\}$. Then*

$$\int_{-R}^R |g_\lambda(x)| dx \leq cR^2 \left(\frac{1}{\sqrt{n}} + \frac{1}{\lambda} \right) + \frac{c' \log(\lambda)}{\lambda}.$$

Proof. Recall that $g_\lambda(x) = g * K_\lambda(x)$, so that

$$(4) \quad \begin{aligned} \int_{-R}^R |g_\lambda(x)| dx &= \frac{1}{\pi} \int_{-R}^R \left| \int_{-\infty}^{\infty} g\left(x - \frac{2w}{\lambda}\right) \left(\frac{\sin(w)}{w}\right)^2 dw \right| dx \\ &\leq \frac{2}{\pi} \int_0^R \int_{-\infty}^{\infty} \min \left\{ \frac{2}{|x - \frac{2w}{\lambda}|}, \frac{C|x - \frac{2w}{\lambda}|}{\sqrt{n}} \right\} \left(\frac{\sin(w)}{w}\right)^2 dw dx. \end{aligned}$$

Assume that λ will be chosen with $\lambda > 2$. For $x > 0$,

$$\begin{aligned} &\int_0^{\infty} \min \left\{ \frac{2}{|x - \frac{2w}{\lambda}|}, \frac{C|x - \frac{2w}{\lambda}|}{\sqrt{n}} \right\} \left(\frac{\sin(w)}{w}\right)^2 dw \\ &\leq \int_0^1 \frac{C|x - \frac{2w}{\lambda}|}{\sqrt{n}} dw + \int_1^{\frac{\lambda}{2}(1+x)} \frac{C|x - \frac{2w}{\lambda}|}{w^2 \sqrt{n}} dw + \int_{\frac{\lambda}{2}(1+x)}^{\infty} \frac{2}{w^2 |x - \frac{2w}{\lambda}|} dw. \end{aligned}$$

Now, the first term is trivially bounded by $\frac{C(x+1)}{\sqrt{n}}$. For the second term,

$$\begin{aligned} &\int_1^{\frac{\lambda}{2}(1+x)} \frac{C|x - \frac{2w}{\lambda}|}{w^2 \sqrt{n}} dw \\ &\leq \frac{C}{\sqrt{n}} \int_1^{\frac{\lambda}{2}(1+x)} \frac{x + \frac{2w}{\lambda}}{w^2} dw = \frac{C}{\sqrt{n}} \left[x \left(1 - \frac{2}{\lambda(1+x)}\right) + \frac{2}{\lambda} \log \left(\frac{\lambda}{2}(1+x)\right) \right]. \end{aligned}$$

For the final term,

$$\int_{\frac{\lambda}{2}(1+x)}^{\infty} \frac{2}{w^2 |x - \frac{2w}{\lambda}|} dw = \frac{4}{\lambda x^2} \int_{1+\frac{1}{x}}^{\infty} \frac{1}{t^2(t-1)} dt = \frac{4}{\lambda x^2} \left(\log(1+x) - \frac{x}{1+x} \right),$$

and so

$$\begin{aligned} &\int_0^R \int_0^{\infty} \min \left\{ \frac{2}{|x - \frac{2w}{\lambda}|}, \frac{C|x - \frac{2w}{\lambda}|}{\sqrt{n}} \right\} \left(\frac{\sin(w)}{w}\right)^2 dw dx \\ &\leq \int_0^R \left(\frac{C}{\sqrt{n}} \left[2x + 1 + \frac{2}{\lambda} \log \left(\frac{\lambda}{2}(1+x)\right) - \frac{2x}{\lambda(1+x)} \right] + \frac{4}{\lambda x^2} \left(\log(1+x) - \frac{x}{1+x} \right) \right) dx \\ &\leq cR^2 \left(\frac{1}{\sqrt{n}} + \frac{1}{\lambda} \right) + \frac{c \log(\lambda)}{\lambda}. \end{aligned}$$

Similarly, for $x < 0$,

$$\begin{aligned} &\int_{-\infty}^0 \min \left\{ \frac{2}{|x - \frac{2w}{\lambda}|}, \frac{C|x - \frac{2w}{\lambda}|}{\sqrt{n}} \right\} \left(\frac{\sin(w)}{w}\right)^2 dw \\ &\leq \int_0^1 \frac{C(x + \frac{2w}{\lambda})}{\sqrt{n}} dw + \int_1^{\frac{\lambda}{2}(1+x)} \frac{C(x + \frac{2w}{\lambda})}{w^2 \sqrt{n}} dw + \int_{\frac{\lambda}{2}(1+x)}^{\infty} \frac{2}{w^2 (x + \frac{2w}{\lambda})} dw \\ &\leq \frac{C(x+1)}{\sqrt{n}} + \frac{C}{\sqrt{n}} \left[x \left(1 - \frac{2}{\lambda(1+x)}\right) + \frac{2}{\lambda} \log \left(\frac{\lambda}{2}(1+x)\right) \right] + \frac{4}{\lambda x^2} \left[\frac{x}{x+1} - \log \left(\frac{2x+1}{x+1}\right) \right], \end{aligned}$$

and so

$$\begin{aligned}
& \int_0^R \int_{-\infty}^0 \min \left\{ \frac{2}{|x - \frac{2w}{\lambda}|}, \frac{C|x - \frac{2w}{\lambda}|}{\sqrt{n}} \right\} \left(\frac{\sin(w)}{w} \right)^2 dw dx \\
& \leq \int_0^R \left(\frac{C}{\sqrt{n}} \left[2x + 1 + \frac{2}{\lambda} \log \left(\frac{\lambda}{2}(1+x) \right) - \frac{2x}{\lambda(1+x)} \right] + \frac{4}{\lambda x^2} \left[\frac{x}{x+1} - \log \left(\frac{2x+1}{x+1} \right) \right] \right) dx \\
& \leq cR^2 \left(\frac{1}{\sqrt{n}} + \frac{1}{\lambda} \right) + \frac{c' \log(\lambda)}{\lambda}.
\end{aligned}$$

□

Proposition 7. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ have $\|f\|_{BL} \leq 1$ and $\text{supp}(f) \subseteq [-2R, 2R]$. Then*

$$\left| \int f_\lambda d\mu_n^{DOS} - \int f_\lambda d\gamma \right| \leq cR^2 \left(\frac{1}{\sqrt{n}} + \frac{1}{\lambda} \right) + \frac{c' \log(\lambda)}{\lambda}.$$

Proof. Let $\psi_n(\xi)$ denote the characteristic function of μ_n^{DOS} and $\varphi(\xi)$ the characteristic function of the standard Gaussian distribution γ . Then by the various definitions,

$$\begin{aligned}
& \left| \int f_\lambda d\mu_n^{DOS} - \int f_\lambda d\gamma \right| \\
& = \frac{1}{2\pi} \left| \int_{-\infty}^{\infty} \int_{-\lambda}^{\lambda} \hat{f}(\xi) e^{ix\xi} \left(1 - \frac{|\xi|}{\lambda} \right) d\xi d\mu_n^{DOS}(x) - \int_{-\infty}^{\infty} \int_{-\lambda}^{\lambda} \hat{f}(\xi) e^{ix\xi} \left(1 - \frac{|\xi|}{\lambda} \right) d\xi d\gamma(x) \right| \\
& = \frac{1}{2\pi} \left| \int_{-\lambda}^{\lambda} \hat{f}(\xi) \left(1 - \frac{|\xi|}{\lambda} \right) \left(\int_{-\infty}^{\infty} e^{ix\xi} d\mu_n^{DOS}(x) - \int_{-\infty}^{\infty} e^{ix\xi} d\gamma(x) \right) d\xi \right| \\
& = \frac{1}{2\pi} \left| \int_{-\lambda}^{\lambda} \hat{f}(\xi) \left(1 - \frac{|\xi|}{\lambda} \right) (\psi_n(-\xi) - \varphi(-\xi)) d\xi \right| \\
& = \frac{1}{2\pi} \left| \int_{-\lambda}^{\lambda} \left(\int_{-\infty}^{\infty} f(x) e^{-ix\xi} dx \right) \left(1 - \frac{|\xi|}{\lambda} \right) (\psi_n(-\xi) - \varphi(-\xi)) d\xi \right| \\
& = \frac{1}{2\pi} \left| \int_{-\lambda}^{\lambda} \left(\int_{-\infty}^{\infty} \frac{f'(x)}{i\xi} e^{-ix\xi} dx \right) \left(1 - \frac{|\xi|}{\lambda} \right) (\psi_n(-\xi) - \varphi(-\xi)) d\xi \right| \\
& = \frac{1}{2\pi} \left| \int_{-2R}^{2R} f'(x) \int_{-\lambda}^{\lambda} \left(\frac{\psi_n(\xi) - \varphi(\xi)}{\xi} \right) \left(1 - \frac{|\xi|}{\lambda} \right) e^{ix\xi} d\xi dx \right| \\
& = \left| \int_{-2R}^{2R} f'(x) \left(\frac{\psi_n(\xi) - \varphi(\xi)}{\xi} \right)_\lambda(x) dx \right|,
\end{aligned}$$

where the third to last line follows by integration by parts and we have used that f is supported on $[-2R, 2R]$ in the last two lines.

We can now apply the result of Lemma 6 and the fact the $\|f\|_{BL} \leq 1$ to obtain the conclusion.

□

We are now ready to give the proof of Theorem 3

Proof of Theorem 3. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ have $\|f\|_{BL} \leq 1$. Then by Lemma 4 (and its much stronger counterpart for γ),

$$\sup_{|f|_L \leq 1} \left| \int f d\mu_n^{DOS} - \int f d\gamma \right| \leq \frac{c}{R^2} + \sup_{\substack{|f|_L \leq 1, \\ \text{supp}(f) \subseteq [-2R, 2R]}} \left| \int f d\mu_n^{DOS} - \int f d\gamma \right|.$$

By Lemma 5, for f with $\|f\|_{BL} \leq 1$ and support in $[-2R, 2R]$,

$$\sup_x |f(x) - f_\lambda(x)| \leq \frac{8 \log(\lambda) + 8 \log(2R) + 6}{\pi \lambda},$$

and so

$$\begin{aligned} & \sup_{\substack{|f|_L \leq 1, \\ \text{supp}(f) \subseteq [-2R, 2R]}} \left| \int f d\mu_n^{DOS} - \int f d\gamma \right| \\ & \leq \sup_{\substack{|f|_L \leq 1, \\ \text{supp}(f) \subseteq [-2R, 2R]}} \left| \int f_\lambda d\mu_n^{DOS} - \int f_\lambda d\gamma \right| + \frac{16 \log(\lambda) + 16 \log(2R) + 12}{\pi \lambda}. \end{aligned}$$

Applying Proposition 7 now gives

$$d_{BL}(\mu_n^{DOS}, \gamma) \leq \frac{c}{R^2} + \frac{16 \log(\lambda) + 16 \log(2R) + 12}{\pi \lambda} + cR^2 \left(\frac{1}{\sqrt{n}} + \frac{1}{\lambda} \right) + \frac{c' \log(\lambda)}{\lambda}.$$

Choosing $\lambda = n$ and $R = n^{1/8}$ completes the proof. \square

3. CONCENTRATION AND AVERAGE DISTANCE TO AVERAGE

A crucial underpinning of the remainder of the proof is the following concentration of measure property of a Gaussian random vector.

Proposition 8 (See, e.g., Ch. 1 of [8]). *Let $(Z_k)_{1 \leq k \leq n}$ be a standard n -dimensional Gaussian random vector, and let $F : \mathbb{R}^n \rightarrow \mathbb{R}$ be Lipschitz with Lipschitz constant L . There are universal constants C, c such that*

$$\mathbb{P}[|F(Z_1, \dots, Z_n) - \mathbb{E}F(Z_1, \dots, Z_n)| > t] \leq Ce^{-ct^2/L^2}.$$

The concentration phenomenon is key both in proving the concentration of the bounded-Lipschitz distance from μ_n to a fixed reference measure, and in estimating $\mathbb{E}c_{BL}(\mu_n, \mu_n^{DOS})$. The following lemma gives the necessary Lipschitz estimates for this approach.

Lemma 9. *Let $\mathbf{x} = \{x_{a,b,j}\} \in \mathbb{R}^{9n}$ (with, say, lexicographic ordering). Define $H_n(\mathbf{x})$ by*

$$H_n(\mathbf{x}) := \frac{1}{3\sqrt{n}} \sum_{a,b=1}^3 \sum_{j=1}^n x_{a,b,j} \sigma_j^{(a)} \sigma_{j+1}^{(b)},$$

and let μ_n be the spectral measure of $H_n(\mathbf{x})$. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ have $\|f\|_{BL} \leq 1$. Then

(a) the map

$$\mathbf{x} \mapsto \int f d\mu_n$$

is $\frac{1}{3\sqrt{n}}$ -Lipschitz, and

(b) for any probability measure ρ on \mathbb{R} , the map

$$\mathbf{x} \mapsto d_{BL}(\mu_n, \rho)$$

is $\frac{1}{3\sqrt{n}}$ -Lipschitz.

Proof. First consider the map $\mathbf{x} \mapsto H_n(\mathbf{x})$, and equip the space of $2^n \times 2^n$ symmetric matrices with the Hilbert-Schmidt norm:

$$\|A\|_{H.S.} := \text{Tr}(AA^T) = \text{Tr}(A^2).$$

For $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^{9n}$, write $H_n := H_n(\mathbf{x})$ and $H'_n := H_n(\mathbf{x}')$. Then

$$\begin{aligned} \|H_n - H'_n\|_{H.S.}^2 &= \text{Tr}[(H_n - H'_n)^2] \\ &= \frac{1}{9n} \sum_{j,k=1}^n \sum_{a,b,c,d=1}^3 (x_{a,b,j} - x'_{a,b,j})(x_{c,d,k} - x'_{c,d,k}) \text{Tr}(\sigma_j^{(a)} \sigma_{j+1}^{(b)} \sigma_k^{(c)} \sigma_{k+1}^{(d)}). \end{aligned}$$

Recall that $\text{Tr}(A \otimes B) = \text{Tr}(A) \text{Tr}(B)$, and that $\text{Tr}(\sigma^{(a)}) = 0$ for each of the Pauli matrices. If $j \neq k$, then the matrix $\sigma_j^{(a)} \sigma_{j+1}^{(b)} \sigma_k^{(c)} \sigma_{k+1}^{(d)}$ is a tensor product, at least two of whose factors are Pauli matrices; that is, if $j \neq k$, then

$$\text{Tr}(\sigma_j^{(a)} \sigma_{j+1}^{(b)} \sigma_k^{(c)} \sigma_{k+1}^{(d)}) = 0.$$

If $j = k$, then

$$\sigma_j^{(a)} \sigma_{j+1}^{(b)} \sigma_k^{(c)} \sigma_{k+1}^{(d)} = I_2^{\otimes(j-1)} \otimes \sigma^{(a)} \sigma^{(c)} \otimes \sigma^{(b)} \sigma^{(d)} \otimes I_2^{\otimes(n-j-1)},$$

and thus

$$\text{Tr}(\sigma_j^{(a)} \sigma_{j+1}^{(b)} \sigma_k^{(c)} \sigma_{k+1}^{(d)}) = \begin{cases} 2^n, & a = c, b = d; \\ 0, & \text{otherwise.} \end{cases}$$

It follows that

$$\|H_n - H'_n\|_{H.S.} = \sqrt{\frac{2^n}{9n} \sum_{j=1}^n \sum_{a,b=1}^3 (x_{a,b,j} - x'_{a,b,j})^2} = \frac{2^{n/2}}{3\sqrt{n}} \|\mathbf{x} - \mathbf{x}'\|,$$

and so the map $\mathbf{x} \mapsto H_n$ is $\frac{2^{n/2}}{3\sqrt{n}}$ -Lipschitz.

Now consider the map $H_n \mapsto \int f d\mu_n$. By definition, $\int f d\mu_n = \frac{1}{2^n} \sum_{j=1}^{2^n} f(\lambda_j)$ so

$$\left| \int f d\mu_n - \int f d\mu'_n \right| = 2^{-n} \left| \sum_{j=1}^{2^n} f(\lambda_j) - f(\lambda'_j) \right| \leq 2^{-n} \sum_{j=1}^{2^n} |\lambda_j - \lambda'_j| \leq 2^{-n/2} \sqrt{\sum_{j=1}^{2^n} |\lambda_j - \lambda'_j|^2},$$

making use of the fact that f is 1-Lipschitz.

The Hoffman-Wielandt inequality (see, e.g., [1, Theorem VI.4.1]) gives that

$$2^{-n/2} \sqrt{\sum_{j=1}^{2^n} |\lambda_j - \lambda'_j|^2} \leq 2^{-n/2} \|H_n - H'_n\|_{H.S.}.$$

and so the map $H_n \mapsto \int f d\mu_n$ is $2^{-n/2}$ -Lipschitz; this completes the proof of part (a).

For part (b), first note that by the triangle inequality for d_{BL} ,

$$|d_{BL}(\mu_n, \rho) - c_{BL}(\mu'_n, \rho)| \leq c_{BL}(\mu_n, \mu'_n).$$

Define a coupling π of μ_n and μ'_n by

$$\pi := \frac{1}{2^n} \sum_{j=1}^{2^n} \delta_{(\lambda_j, \lambda'_j)},$$

where λ_j and λ'_j are ordered eigenvalues of H_n and H'_n respectively. Then by the Kantorovich–Rubinstein theorem,

$$d_{BL}(\mu_n, \mu'_n) \leq W_1(\mu_n, \mu'_n) \leq \int |x - y| d\pi(x, y) = \frac{1}{2^n} \sum_j |\lambda_j - \lambda'_j|.$$

Applying the Cauchy-Schwarz inequality and the Hoffman-Wielandt inequality exactly as before gives that the map $H_n \mapsto d_{BL}(\mu_n, \rho)$ is $2^{-n/2}$ -Lipschitz; together with the Lipschitz estimate for $\mathbf{x} \mapsto H_n(\mathbf{x})$ given above, this completes the proof. \square

It is thus immediate from Proposition 8 and Lemma 9 that if ρ is any probability measure,

$$(5) \quad \mathbb{P}[|d_{BL}(\mu_n, \rho) - \mathbb{E}d_{BL}(\mu_n, \rho)| > t] \leq Ce^{-cnt^2},$$

and so part (b) of Theorem 1 follows immediately from part (a).

To prove part (a), recall that in the previous section it was shown that

$$d_{BL}(\mu_n^{DOS}, \gamma) \leq \frac{C}{n^{1/4}}.$$

It thus suffices by the triangle inequality for d_{BL} to show that

$$(6) \quad \mathbb{E}d_{BL}(\mu_n, \mu_n^{DOS}) = \sup_{\|f\|_{BL} \leq 1} \left(\int f d\mu_n - \int f d\mu_n^{DOS} \right) \leq \frac{C}{n^{1/6}}.$$

Observe first that f with $\|f\|_{BL} \leq 1$, if

$$X_f := \int f d\mu_n - \int f d\mu_n^{DOS},$$

then $\mathbb{E}X_f = 0$ and by Proposition 8 and Lemma 9,

$$\mathbb{P}[|X_f| > t] \leq Ce^{-cnt^2}.$$

More generally,

$$(7) \quad \mathbb{P}[|X_f - X_g| > t] = \mathbb{P}[|X_{f-g}| > t] \leq Ce^{-\frac{cnt^2}{\|f-g\|_{BL}^2}};$$

that is, the process $\{X_f\}$ indexed by functions $f : \mathbb{R} \rightarrow \mathbb{R}$ with $\|f\|_{BL} \leq 1$ and $f(0) = 0$ is a sub-Gaussian stochastic process, with respect to the distance $d(f, g) = \frac{\|f-g\|_{BL}}{\sqrt{n}}$.

The idea at this point is to use Dudley's entropy bound to estimate the expected supremum of this process, but to do this successfully, a series of approximations must be made first to reduce the size of the (currently infinite-dimensional) indexing set of the process. The first step is a somewhat more sophisticated truncation argument than the one which appeared in Section 2, which allows us to assume that our test functions are finitely supported.

Let $\|A\|_{op}$ denote the $\ell_2 \rightarrow \ell_2$ operator norm of a matrix A , and observe that the map $\mathbf{x} \mapsto \|H_n\|_{op}$ is 1-Lipschitz:

$$\begin{aligned} |\|H_n\|_{op} - \|H'_n\|_{op}| &\leq \|H_n - H'_n\|_{op} \\ &= \frac{1}{3\sqrt{n}} \left\| \sum_{j=1}^n \sum_{a,b=1}^3 (x_{a,b,j} - x'_{a,b,j}) \sigma_j^{(a)} \sigma_{j+1}^{(b)} \right\|_{op} \\ &\leq \frac{1}{3\sqrt{n}} \sum_{j=1}^n \sum_{a,b=1}^3 |(x_{a,b,j} - x'_{a,b,j})| \|\sigma_j^{(a)} \sigma_{j+1}^{(b)}\|_{op}. \end{aligned}$$

Now, $\|A \otimes B\|_{op} = \|A\|_{op} \|B\|_{op}$ and all of the Pauli matrices have operator norm 1, so $\|\sigma_j^{(a)} \sigma_{j+1}^{(b)}\|_{op} = 1$ for all a, b, j . An application of the Cauchy-Schwarz inequality thus gives that

$$|\|H_n\|_{op} - \|H'_n\|_{op}| \leq \sqrt{\sum_{j=1}^n \sum_{a,b=1}^3 |x_{a,b,j} - x'_{a,b,j}|^2} = \|\mathbf{x} - \mathbf{x}'\|.$$

It thus follows from concentration of measure that

$$\mathbb{P}[|\|H_n\|_{op} - \mathbb{E}\|H_n\|_{op}| > t] \leq Ce^{-ct^2}.$$

Since

$$\mathbb{E}\|H_n\|_{op} \leq \frac{1}{3\sqrt{n}} \sum_{a,b=1}^3 \sum_{j=1}^n \mathbb{E}|Z_{a,b,j}| = \frac{3\sqrt{2n}}{\sqrt{\pi}}.$$

one has in particular that if $t \geq C\sqrt{n}$, then

$$\mathbb{P}[\|H_n\|_{op} > t] \leq Ce^{-ct^2}.$$

One can interpret this statement as saying that if $R \sim \sqrt{n}$ it is extremely unlikely that H_n will have any eigenvalues outside $[-R, R]$, and so truncation of test functions to that interval should not result in much loss.

More specifically, if f_R is the truncation of f to $[-2R, 2R]$ given in Equation (3) of Section 2, then

$$\mathbb{E} \left| \int (f - f_R) d\mu_n \right| \leq C\mu_n^{DOS}(\{x : |x| > R\}).$$

Since

$$\mu_n^{DOS}(\{x : |x| > R\}) = \frac{1}{2n} \mathbb{E} |\{j : |\lambda_j| > R\}| \leq \mathbb{P}[\|H_n\|_{op} > R],$$

it follows that

$$\mathbb{E} \left| \int (f - f_R) d\mu_n \right| \leq Ce^{-cnR^2}.$$

The indexing space of the process $\{X_f\}$ may thus be safely reduced to those f supported on $[-2R, 2R]$ with R of order \sqrt{n} , with an error which is exponentially small in n .

The next step in reducing the indexing space is to approximate bounded Lipschitz test functions by piecewise linear ones. Given $f : \mathbb{R} \rightarrow \mathbb{R}$ with $\|f\|_{BL} \leq 1$ and $\text{supp}(f) \subseteq [-2R, 2R]$, consider the piecewise linear function $g : \mathbb{R} \rightarrow \mathbb{R}$ defined so that $\text{supp}(g) \subseteq$

$[-2R, 2R]$, and $g(x) = f(x)$ at each x of the form $-2R + \frac{4Rk}{m}$, for $0 \leq k \leq m$. Because f is 1-Lipschitz,

$$\|f - g\|_\infty \leq \frac{2R}{m},$$

and so

$$|X_f - X_g| \leq \frac{4R}{m}.$$

It follows that

$$(8) \quad \mathbb{E} d_{BL}(\mu_n, \mu_n^{DOS}) \leq \mathbb{E} \sup_{g \in \mathcal{G}} X_g + \frac{4R}{m} + Ce^{-cR^2},$$

where the supremum is taken over the class \mathcal{G} of functions $g : \mathbb{R} \rightarrow \mathbb{R}$ satisfying

- $g(0) = 0$;
- $\|g\|_{BL} \leq 1$;
- $\text{supp}(g) \subseteq [-2R, 2R]$;
- g is linear on intervals of the form $\left[-2R + \frac{4Rk}{m}, -2R + \frac{4R(k+1)}{m}\right]$.

With the reduction to \mathcal{G} as the indexing space of our stochastic process, it is now possible to apply Dudley's entropy bound (see, e.g., the introduction of [11]):

Proposition 10 (Dudley). *Let $\{Y_x : x \in M\}$ be a centered subgaussian stochastic process indexed by the metric space (M, d) . Then*

$$\mathbb{E} \sup_{x \in M} |Y_x| \leq K \int_0^\infty \sqrt{\log N(M, d, \epsilon)} d\epsilon,$$

where $N(M, d, \epsilon)$ denotes the number of ϵ -balls (with respect to the metric d) needed to cover M , and $K > 0$ depends only on the constants of the sub-Gaussian increment condition.

Let \mathcal{G} denote the index set described above. Applying Proposition 10 to $\{X_g\}_{g \in \mathcal{G}}$ gives that

$$\mathbb{E} \sup_{g \in \mathcal{G}} X_g \leq K \int_0^\infty \sqrt{\log \left[N \left(\mathcal{G}, \frac{|\cdot|_L}{\sqrt{n}}, \epsilon \right) \right]} d\epsilon = \frac{K}{\sqrt{n}} \int_0^\infty \sqrt{\log [N(\mathcal{G}, |\cdot|_L, \epsilon)]} d\epsilon.$$

Since $(\mathcal{G}, |\cdot|_L)$ is just an $(m+1)$ -dimensional normed space, standard volumetric estimates (see [10, Lemma 2.6]) give that

$$N(\mathcal{G}, |\cdot|_L, \epsilon) \leq \left(\frac{3}{\epsilon} \right)^{m+1},$$

so that together with (8), we have that

$$\mathbb{E} W_1(\mu_{H_n}, \mu_n^{DOS}) \leq K \sqrt{\frac{m}{n}} + \frac{2R}{m} + Ce^{-cR^2}.$$

Choosing R of order \sqrt{n} and $m = n^{2/3}$ completes the proof of (6), and thus the proof of part (a) of Theorem 1.

4. THE SPHERICAL MODEL

As discussed in the introduction, all of the random matrix models of quantum spin chains considered so far involve independent coefficients. A perhaps more geometrically natural alternative is to consider the random matrix

$$(9) \quad H_n := \sum_{j=1}^n \sum_{a,b=1}^3 x_{a,b,j} \sigma_j^{(a)} \sigma_{j+1}^{(b)},$$

where the $\sigma_j^{(a)}$ are as before, but the vector of coefficients $\mathbf{x} = \{x_{a,b,j}\}_{\substack{1 \leq a,b \leq 3 \\ 1 \leq j \leq n}}$ is chosen uniformly from the unit sphere in \mathbb{R}^{9n} . While this model introduces dependence among the coefficients, it is still possible to prove the almost sure convergence of the empirical spectral measure μ_n of H_n to the standard Gaussian distribution, albeit without a specific rate.

Theorem 11. *For each $n \geq 1$, let μ_n be the spectral measure of the random matrix H_n defined as in (9). Then almost surely, the sequence $\{\mu_n\}_{n \in \mathbb{N}}$ tends weakly to the standard Gaussian distribution, as $n \rightarrow \infty$.*

It should be noted that the statement above implicitly assumes some joint distribution of the coefficient vectors, but the theorem is true independent of what that joint distribution is.

That the earlier results for i.i.d. coefficients can be extended to this dependent setting relies on two important properties of uniform random vectors on the sphere. The first is that explicit computations are still at least somewhat feasible. The second is that the concentration of measure phenomenon which was strongly used in the Gaussian case holds for random vectors on the sphere as well, as follows.

Lemma 12 (Lévy's lemma; see [Ch. 1 of [8]]). *Let $(x_k)_{1 \leq k \leq n}$ be a random vector, uniformly distributed on \mathbb{S}^{n-1} , and let $F : \mathbb{R}^n \rightarrow \mathbb{R}$ be Lipschitz with Lipschitz constant L . There are universal constants C, c such that*

$$\mathbb{P}[|F(x_1, \dots, x_n) - \mathbb{E}F(x_1, \dots, x_n)| > t] \leq Ce^{-cnt^2/L^2}.$$

The following modified version of Theorem 1 holds for the spherical model; here we compare μ_n to μ_n^{DOS} rather than to the Gaussian distribution.

Theorem 13. *Let μ_n be the spectral measure of H_n and let $\mu_n^{DOS} := \mathbb{E}\mu_n$ be the density of states measure. There are universal constants C, C' and c such that*

$$(a) \quad \mathbb{E}d_{BL}(\mu_n, \mu_n^{DOS}) \leq \frac{C}{n^{1/6}};$$

$$(b) \quad \mathbb{P}\left[d_{BL}(\mu_n, \mu_n^{DOS}) \geq \frac{C}{n^{1/6}} + t\right] \leq Ce^{-cnt^2};$$

and

(c) *with probability 1, for all sufficiently large n ,*

$$d_{BL}(\mu_n, \mu_n^{DOS}) \leq \frac{C'}{n^{1/6}}.$$

All of the proofs in Section 3 go through in exactly the same way, using Lévy's lemma in place of Proposition 8. (Note the difference in normalization: Proposition 8 is stated for a standard Gaussian random vector, with expected length on the order of \sqrt{n} , whereas Lévy's lemma is stated for a random vector on the unit sphere.) The missing element in

showing the almost sure convergence of μ_n to the Gaussian distribution is the comparison of μ_n^{DOS} to Gaussian, which in the case of i.i.d. Gaussian coefficients in H_n followed from the characteristic function estimate proved in [7]. The following result gives an analog of their result in for the spherical model, but without a similarly good rate of convergence; this is the reason that we do not obtain an almost sure convergence rate of μ_n in the spherical model.

Proposition 14. *Let $\psi_n(t)$ denote the characteristic function of the density of states measure of H_n as defined in Equation (9). Then for each $t \in \mathbb{R}$,*

$$\psi_n(t) \xrightarrow{n \rightarrow \infty} e^{-\frac{t^2}{2}}.$$

To modify the approach in [7] to prove Proposition 14, we will need to calculate expectations of certain functions over the unit sphere; the following lemma gives explicit formulae.

Lemma 15 (See [4]). *Let $P(x) = |x_1|^{\alpha_1} |x_2|^{\alpha_2} \cdots |x_n|^{\alpha_n}$. Then if X is uniformly distributed on S^{n-1} ,*

$$\mathbb{E}[P(X)] = \frac{\Gamma(\beta_1) \cdots \Gamma(\beta_n) \Gamma(\frac{n}{2})}{\Gamma(\beta_1 + \cdots + \beta_n) \pi^{n/2}},$$

where $\beta_i = \frac{1}{2}(\alpha_i + 1)$ for $1 \leq i \leq n$ and

$$\Gamma(t) = \int_0^\infty s^{t-1} e^{-s} ds = 2 \int_0^\infty r^{2t-1} e^{-r^2} dr.$$

The crucial technical ingredient for Proposition 14 is the following.

Lemma 16. *Let $\{x_k\}_{1 \leq k \leq N}$ be uniformly distributed on the unit sphere in \mathbb{R}^N . For each $t \in \mathbb{R}$,*

$$\mathbb{E} \left[\prod_{k=1}^N \cos(tx_k) \right] \xrightarrow{N \rightarrow \infty} e^{-t^2}.$$

Proof. We first show that it suffices to approximate the cosine with a second-order Taylor expansion. It follows from the Lévy's lemma and the fact that $|\cos(tx_k)| \leq 1$ that

$$\begin{aligned} \left| \mathbb{E} \prod_{k=1}^N \cos(tx_k) - \mathbb{E} \prod_{k=1}^N \cos(tx_k) \mathbb{1}_{\left|1 - \frac{(tx_k)^2}{2}\right| \leq 1} \right| &= \left| \mathbb{E} \left[\prod_{k=1}^N \cos(tx_k) \mathbb{1}_{\bigcup_{k=1}^N \left\{ \left|1 - \frac{(tx_k)^2}{2}\right| > 1 \right\}} \right] \right| \\ (10) \quad &\leq \mathbb{P} \left[\bigcup_{k=1}^N \left\{ \left|1 - \frac{(tx_k)^2}{2}\right| > 1 \right\} \right] \\ &= \mathbb{P} \left[\bigcup_{k=1}^N \left\{ \frac{(tx_k)^2}{2} > 2 \right\} \right] \leq CN e^{-\frac{cN}{t^2}}. \end{aligned}$$

From the trivial estimate that $|z_1 \cdots z_m - w_1 \cdots w_m| \leq \sum_{k=1}^m |z_k - w_k|$ if $|z_k|, |w_k| \leq 1$ for all k , it then follows that

$$\begin{aligned} (11) \quad \left| \mathbb{E} \prod_{k=1}^N \left\{ \cos(tx_k) - 1 + \frac{(tx_k)^2}{2} \right\} \mathbb{1}_{\left|1 - \frac{(tx_k)^2}{2}\right| \leq 1} \right| &\leq \sum_{k=1}^N \mathbb{E} \left| \cos(tx_k) - 1 + \frac{(tx_k)^2}{2} \right| \mathbb{1}_{\left|1 - \frac{(tx_k)^2}{2}\right| \leq 1} \\ &\leq \frac{N \mathbb{E}(tx_1)^4}{4!} \leq \frac{t^4}{N \cdot 4!}, \end{aligned}$$

where the last line follows from Lemma 15. Now, by the Cauchy-Schwarz inequality,

$$\begin{aligned}
 (12) \quad & \left| \mathbb{E} \prod_{k=1}^N \left\{ 1 - \frac{(tx_k)^2}{2} \right\} \mathbb{1}_{\left| 1 - \frac{(tx_k)^2}{2} \right| \leq 1} - \mathbb{E} \prod_{k=1}^N \left\{ 1 - \frac{(tx_k)^2}{2} \right\} \right| \\
 & \leq \sqrt{\mathbb{E} \prod_{k=1}^N \left(1 - \frac{(tx_k)^2}{2} \right)^2} \sqrt{\mathbb{P} \left[\bigcup_{k=1}^N \left\{ \left| 1 - \frac{(tx_k)^2}{2} \right| > 1 \right\} \right]} \\
 & \leq C e^{-\frac{cN}{t^2}} \sqrt{\mathbb{E} \prod_{k=1}^N \left(1 - (tx_k)^2 + \frac{(tx_k)^4}{4} \right)}.
 \end{aligned}$$

Expanding this last expression and using Lemma 15 gives that

$$\begin{aligned}
 (13) \quad & \mathbb{E} \prod_{k=1}^N \left(1 - (tx_k)^2 + \frac{(tx_k)^4}{4} \right) = \sum_{j=0}^N \binom{N}{j} \sum_{\ell=0}^{N-j} \binom{N-j}{\ell} \frac{(-t)^{2j} t^{4\ell}}{4^\ell} \mathbb{E}[x_1^2 \cdots x_j^2 x_{j+1}^4 \cdots x_{j+\ell}^4] \\
 & \leq \sum_{j=0}^N \binom{N}{j} \left(\frac{-t^2}{N} \right)^j \sum_{\ell=0}^{N-j} \binom{N-j}{\ell} \left(\frac{3t^4}{4N^2} \right)^\ell \\
 & = \sum_{j=0}^N \binom{N}{j} \left(\frac{-t^2}{N} \right)^j \left(1 + \frac{3t^4}{4N^2} \right)^{N-j} \\
 & = \left(1 - \frac{t^2}{N} + \frac{3t^4}{4N^2} \right)^N.
 \end{aligned}$$

Since this last expression is asymptotic to e^{-t^2} , it is in particular bounded. Combining equations (10), (11), (12), and (13) gives that

$$(14) \quad \left| \mathbb{E} \prod_{k=1}^N \cos(tx_k) - \mathbb{E} \prod_{k=1}^N \left(1 - \frac{(tx_k)^2}{2} \right) \right| \leq \frac{Ct^4}{N},$$

and it remains to analyze $\mathbb{E} \prod_{k=1}^N \left(1 - \frac{(tx_k)^2}{2} \right)$.

By Lemma 15,

$$\begin{aligned}
 \mathbb{E} \left[\prod_{k=1}^N \left(1 - \frac{(tx_k)^2}{2} \right) \right] &= \sum_{k=0}^N \binom{N}{k} \left(\frac{-t^2}{2} \right)^k \frac{\Gamma(\frac{N}{2})}{2^k \Gamma(\frac{N}{2} + k)} \\
 &= \sum_{k=0}^N \frac{(\frac{-t^2}{2})^k}{k!} \frac{N(N-1) \cdots (N-k+1)}{(N+2k-2)(N+2k-4) \cdots (N)} \\
 &= \sum_{k=0}^N \frac{(\frac{-t^2}{2})^k}{k!} \left[\frac{\prod_{\ell=1}^{k-1} (1 - \frac{\ell}{N})}{\prod_{\ell=1}^{k-1} (1 + \frac{2\ell}{N})} \right].
 \end{aligned}$$

Clearly

$$\frac{\prod_{\ell=1}^{k-1} (1 - \frac{\ell}{N})}{\prod_{\ell=1}^{k-1} (1 + \frac{2\ell}{N})} \leq 1,$$

and applying Taylor's theorem to the logarithms gives that there is a constant C such that

$$\begin{aligned} \frac{\prod_{\ell=1}^{k-1} \left(1 - \frac{\ell}{N}\right)}{\prod_{\ell=1}^{k-1} \left(1 + \frac{2\ell}{N}\right)} &= \exp \left(\sum_{\ell=1}^{k-1} \log \left(1 - \frac{\ell}{N}\right) - \log \left(1 + \frac{2\ell}{N}\right) \right) \\ &\geq \exp \left(- \sum_{\ell=1}^{k-1} \left(\frac{3\ell}{N} + \frac{C\ell^2}{N^2} \right) \right) \\ &\geq \exp \left(- \frac{3k(k-1)}{2N} - \frac{Ck^3}{N^2} \right). \end{aligned}$$

It follows that for any $m \leq N$,

$$\begin{aligned} \left| \mathbb{E} \prod_{k=1}^N \left(1 - \frac{(tx_k)^2}{2} \right) - e^{-\frac{t^2}{2}} \right| &\leq \sum_{k=0}^m \frac{\left(\frac{t^2}{2}\right)^k}{k!} \left[1 - \exp \left(- \frac{3k(k-1)}{2N} - \frac{Ck^3}{N^2} \right) \right] + \sum_{k=m+1}^{\infty} \frac{\left(\frac{t^2}{2}\right)^k}{k!} \\ &\leq e^{\frac{t^2}{2}} \left(\frac{3m(m-1)}{2N} + \frac{Cm^3}{N^2} \right) + \frac{\left(\frac{t^2}{2}\right)^{m+1}}{(m+1)!} \\ &\leq \frac{Ce^{\frac{t^2}{2}} m^2}{N} + \frac{C}{\sqrt{m}} \left(\frac{et^2}{2(m+1)} \right)^{m+1}. \end{aligned}$$

Choosing, say, $m = \lceil N^{1/4} \rceil$ completes the proof. \square

Proof of Proposition 14. The proof is a straightforward modification of the one in [7], making use of Lemma 16 instead of the corresponding computation for i.i.d. Gaussian coefficients; below are the details for the necessary modifications, with the part of the proof which is identical to that of [7] omitted.

Suppose that n is even, and make the definitions

$$\begin{aligned} A &:= \sum_{\substack{j=1 \\ j \text{ even}}}^n \sum_{a,b=1}^3 x_{a,b,j} \sigma_j^{(a)} \sigma_{j+1}^{(b)} & B &:= \sum_{\substack{j=1 \\ j \text{ odd}}}^n \sum_{a,b=1}^3 x_{a,b,j} \sigma_j^{(a)} \sigma_{j+1}^{(b)} \\ A_{3(b-1)+a} &:= \sum_{\substack{j=1 \\ j \text{ even}}}^n x_{a,b,j} \sigma_j^{(a)} \sigma_{j+1}^{(b)} & B_{3(b-1)+a} &:= \sum_{\substack{j=1 \\ j \text{ odd}}}^n x_{a,b,j} \sigma_j^{(a)} \sigma_{j+1}^{(b)}. \end{aligned}$$

Then the terms within each sum of each of the A_k and B_k commute, and

$$H_n = A + B = \sum_{k=1}^9 (A_k + B_k).$$

Define

$$\phi_n(t) := \mathbb{E} \left[\frac{1}{2^n} \text{Tr} \left(\prod_{k=1}^9 e^{itA_k} e^{itB_k} \right) \right].$$

Observe that since all of the terms within the A_k and B_k commute,

$$e^{itA_k} = \prod_{\substack{j=1 \\ j \text{ even}}}^n e^{itx_{a,b,j} \sigma_j^{(a)} \sigma_{j+1}^{(b)}} \quad e^{itB_k} = \prod_{\substack{j=1 \\ j \text{ odd}}}^n e^{itx_{a,b,j} \sigma_j^{(a)} \sigma_{j+1}^{(b)}},$$

where $k = 3(b-1) + a$. Now, since the square of any of the Pauli matrices is the identity, it follows from the definition of the matrix exponential in terms of power series that

$$e^{itx_{a,b,j}\sigma_j^{(a)}\sigma_{j+1}^{(b)}} = \cos(tx_{a,b,j})I_{2^n} + \sin(tx_{a,b,j})\sigma_j^{(a)}\sigma_{j+1}^{(b)}.$$

By the symmetry of the uniform distribution on \mathbb{S}^{9n-1} , any term with sine factors in the expansion of $\phi_n(t)$ has vanishing expectation, and so by Lemma 16,

$$\phi_n(t) = \mathbb{E} \left[\prod_{j=1}^n \prod_{a,b=1}^3 \cos(tx_k) \right] \xrightarrow{n \rightarrow \infty} e^{-\frac{t^2}{2}}.$$

At this point the proof can be completed essentially identically to the proof in [7]. \square

From Proposition 14, it follows that μ_n^{DOS} converges weakly to the standard Gaussian distribution; since the bounded-Lipschitz distance is a metric for weak convergence, this means that

$$\lim_{n \rightarrow \infty} d_{BL}(\mu_n^{DOS}, \gamma) = 0.$$

It follows from part (c) of Theorem 13 that

$$\lim_{n \rightarrow \infty} d_{BL}(\mu_n, \mu_n^{DOS}) = 0$$

almost surely, and so Theorem 11 follows.

5. CONCLUDING REMARKS

1. In [7], the authors consider more general distributional assumptions on the coefficients $\{Z_{a,b,j}\}$. Specifically, they consider the random matrices

$$H_n := \sum_{a,b=1}^3 \sum_{j=1}^n \alpha_{a,b,j} \sigma_j^{(a)} \sigma_{j+1}^{(b)},$$

where the $\{\alpha_{a,b,j}\}$ are assumed to be independent and symmetric about 0, with

$$(15) \quad \sum_{a,b=1}^3 \sum_{j=1}^n \mathbb{E} \alpha_{a,b,j}^2 = 1 \quad \lim_{n \rightarrow \infty} \sum_{a,b=1}^3 \sum_{j=1}^n \mathbb{E} |\alpha_{a,b,j}|^{2+\delta} = 0 \quad \max_{a,b,j} \mathbb{E} |\alpha_{a,b,j}|^2 = o\left(\frac{1}{\sqrt{n}}\right),$$

for some $\delta > 0$. The point is that these are the conditions on the $\alpha_{a,b,j}$ under which the Lyapounov central limit holds, which allows the authors to show that the point-wise difference between the characteristic function of μ_n^{DOS} and that of the Gaussian distribution still tends to zero. In order to obtain the same rates of convergence as in the Gaussian case, slightly stronger assumptions on the rate of growth of moments are needed; however, without further assumptions, the concentration arguments used to move to almost sure convergence need not apply.

A probability measure ν is said to satisfy a quadratic transportation cost inequality with constant a if for all probability measures μ absolutely continuous with respect to ν ,

$$(16) \quad W_2(\mu, \nu) \leq \sqrt{aH(\mu||\nu)},$$

where W_2 is the L_2 -Kantorovich distance and $H(\mu||\nu)$ is the relative entropy (or Kullback-Leibler divergence) of μ with respect to ν . If instead of (15), one assumes that the

distribution of each $\alpha_{a,b,j}$ satisfies a quadratic transportation cost inequality with the same constant a , then one can carry out the entire program used here in the Gaussian case with essentially no modification, and one again obtains the almost sure convergence of the spectral measure to Gaussian, with a rate of $\frac{1}{n^{1/6}}$ in W_1 -distance. The assumption that the coefficients $\alpha_{a,b,j}$ all share this property is the most general setting of independent $\alpha_{a,b,j}$ in which the arguments using concentration of measure can be carried out (see [5] for a detailed discussion).

2. In [3], Erdős and Schröder introduced a model for quantum spin glasses with arbitrary coupling geometry. Given a sequence of undirected graphs Γ_n on the vertex sets $\{1, \dots, n\}$ they considered Hermitian random matrices defined by

$$H_n^{\Gamma_n}(\mathbf{x}) = \frac{1}{3\sqrt{e(\Gamma_n)}} \sum_{(ij) \in \Gamma_n} \sum_{a,b=1}^3 \alpha_{a,b,(ij)} \sigma_i^{(a)} \sigma_j^{(b)},$$

where the $\alpha_{a,b,(ij)}$ are assumed to be independent centered random variables with unit variance, and $e(\Gamma_n)$ denotes the number of edges in Γ_n . They proved weak convergence of the density of states measure for this model to the standard Gaussian distribution whenever the maximal degree of a vertex in Γ_n is negligible in comparison to $e(\Gamma_n)$.

A tail bound similar to Theorem 1, part (b) can be obtained for this model if the coefficients are assumed to be standard Gaussian random variables. Using a proof identical to the proof of Lemma 9, one can show that for $\mathbf{x} = \{x_{a,b,(ij)} \in \mathbb{R}^{9e(\Gamma_n)}\}$, the map

$$\mathbf{x} \rightarrow d_{BL}(\mu_n, \rho)$$

is $\frac{1}{3\sqrt{e(\Gamma_n)}}$ -Lipschitz for any probability measure ρ . It then follows from the concentration of measure for standard Gaussian random variables that

$$\mathbb{P}[|d_{BL}(\mu_n, \rho) - \mathbb{E}d_{BL}(\mu_n, \rho)| > t] \leq Ce^{-ce(\Gamma_n)t^2}.$$

Applying this estimate in particular when ρ is the density of states measure, the almost sure convergence of μ_n to the standard Gaussian distribution follows from the Borel-Cantelli Lemma and the convergence proved in [3].

Erdős and Schröder also described a model for the Hamiltonian of a quantum p -spin glasses:

$$H_n^{(p_n\text{-glass})} = 3^{-p_n/2} \binom{n}{p_n}^{-1/2} \sum_{1 \leq i_1 < \dots < i_{p_n} \leq n} \sum_{a_1, \dots, a_{p_n}=1}^3 \alpha_{a_1, \dots, a_{p_n}, (i_1 \dots i_{p_n})} \sigma_{i_1}^{(a_1)} \dots \sigma_{i_{p_n}}^{(a_{p_n})}.$$

They found a sharp phase transition at the threshold $p = \sqrt{n}$ between the standard Gaussian distribution and the Wigner semicircle law, with an explicitly described limiting measure at criticality.

Using the same arguments as above, it can be shown that for $\mathbf{x} = \{x_{a,b,(ij)} \in \mathbb{R}^{9e(\Gamma_n)}\}$, the map

$$\mathbf{x} \rightarrow d_{BL}(\mu_n, \rho)$$

is $3^{-p/2} \binom{n}{p}^{-1/2}$ -Lipschitz for any probability measure ρ . Thus, if $p_n \ll \sqrt{n}$ the empirical spectral measure converges almost surely to the standard Gaussian distribution and if $p_n \gg \sqrt{n}$ the empirical spectral measure converges almost surely to the semicircle law, while if $\frac{p_n}{\sqrt{n}} \xrightarrow{n \rightarrow \infty} \lambda \in (0, \infty)$, the empirical spectral measure converges almost surely to the limiting measure described in [3].

ACKNOWLEDGEMENTS

This research was supported by grant DMS-1308725 from the National Science Foundation. Some of the work was carried out while the second-named author was visiting the Institut de Mathématiques de Toulouse at the Université Paul Sabatier.

REFERENCES

- [1] Rajendra Bhatia. *Matrix analysis*, volume 169 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997.
- [2] Patrick Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, third edition, 1995. A Wiley-Interscience Publication.
- [3] László Erdős and Dominik Schröder. Phase transition in the density of states of quantum spin glasses. *Math. Phys. Anal. Geom.*, 17(3-4):441–464, 2014.
- [4] Gerald B. Folland. How to integrate a polynomial over a sphere. *Amer. Math. Monthly*, 108(5):446–448, 2001.
- [5] Nathael Gozlan. A characterization of dimension free concentration in terms of transportation inequalities. *Ann. Probab.*, 37(6):2480–2498, 2009.
- [6] J. P. Keating, N. Linden, and H. J. Wells. Spectra and eigenstates of spin chain Hamiltonians. *Comm. Math. Phys.*, 338(1):81–102, 2015.
- [7] J. P. Keating, N. Linden, and H. J. Wells. Random matrices and quantum spin chains. *Markov Process. Related Fields*, to appear.
- [8] Michel Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.
- [9] Elizabeth S. Meckes and Mark W. Meckes. Concentration and convergence rates for spectral measures of random matrices. *Probab. Theory Related Fields*, 156(1-2):145–164, 2013.
- [10] Vitali D. Milman and Gideon Schechtman. *Asymptotic theory of finite-dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. With an appendix by M. Gromov.
- [11] Michel Talagrand. *The generic chaining*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2005. Upper and lower bounds of stochastic processes.

DEPARTMENT OF MATHEMATICS, CASE WESTERN RESERVE UNIVERSITY, 10900 EUCLID AVE., CLEVELAND, OHIO 44106, U.S.A.

E-mail address: dab197@case.edu

DEPARTMENT OF MATHEMATICS, CASE WESTERN RESERVE UNIVERSITY, 10900 EUCLID AVE., CLEVELAND, OHIO 44106, U.S.A.

E-mail address: elizabeth.meckes@case.edu